

Figure 1

*Anopheles gambiae* arrestin 1 cDNA sequence (SEQ ID NO: 1)

5  
ACAGGAACGACGGTTGTGATCCCTCCACTGGTGGTGACACGAATCATAAGCATTATTTTCATACCT  
AAAAAACAAAATCTACAAAAAAAGCTTCATTCCCATCGAAAAAAGCTTTCTTGTGAAATCAACCG  
AGCTAACAAACAACATCCTGTGCAAAATCTAGCAGTGAAAGTGTGATATCGTATACCTGTACCTG  
10 TAAACCGTTGTGCGCGTGTGTGCCTTTGTGTATCAATTTTGTGGAAAAACAGAAAATACATCAAAA  
TGTTTACAATTTCAAAGTCTTCAAGAAGTGCGCCCTAATGGAAAGGTTACGCTGTACATGGG  
CAAGCGTGACTTTGTAGACCACGTTTCCGGCGTTGAACCGATCGATGGTATCGTCGTCCTCGAT  
GATGAGTACATTCTGTGACAACCGTAAGGTATTCCGGTCAGATTGTCTGCAGTTTCCGCTACGGCC  
GCGAAGAGGACGAGGTGATGGGACTAACTTCCAGAAGGAGTTATGCCTCGCTTCCGAACAGAT  
CTACCCGCGTCCGGAAAAAGTCGGACAAGGAGCAGACCAAGCTCCAGGAGCGACTGCTGAAGAA  
15 GCTGGGTTCCGAACGCCATCCCGTTCACGTTCAACATCTCGCCGAATGCTCCGTCTTCGGTCACG  
CTGCAGCAGGGCGAAGATGATAATGGAGACCCGTGCGGTGTGTCTGTAAGATCTTTG  
CCGGTGAGTCGGAAACCGATCGTACGCACCGTCGCAGCACCCTACGCTCGGCATACGCAAGAT  
CCAGTTCGCACCGACCAAGCAGGGCCAGCAGCCGTGCACGCTGGTGCAGCAAGGACTTTATGCTA  
AGCCCGGGAGAGCTGGAGCTCGAGGTCACACTAGACAAGCAGCTGTACCTGCACGGGGAGCGA  
20 ATAGGCGTCAACATCTGCATCCGCAACAACCTCGAACAATAATGGTCAAGAAGATTAAGGCCATGG  
TCCAGCAGGGTGTGGATGTGGTGTCTGTTCCAGAATGGTAGCTACCGCAACACAGTGGCATCGCT  
GGAGACTAGCGAGGGTTGCCCAATTGAGCCCGGCTCCAGTCTGCAGAAGGTAATGTACCTCACG  
CCGCTGCTGTCTCTCGAACAAGCAGCGACGTGGCATCGCCCTGGACGGTCAGATCAAGCGTCAGG  
ATCAGTGTGTTGGCCTCGACAACCCCTCTTGGCTCAACCGGATCAGCGAGATGCTTTTCGGCGTTAT  
25 CATATCGTATGCCGTAAAGGTTAAGCTTTTCCTCGGCGCACTCGGCGGCGAGCTGTGCGCGGAA  
CTTCCATTTGTGCTGATGCACCCAAAGCCCGGCACCAAGGCTAAGGTCATCCATGCCGACAGCC  
AGGCCGACGTAGAACTTTCCGACAGGATACAATCGACCGAGGCATGAGTACCTGACTTTGAATA  
GACGACGCAACGGTTTGGAAATGCTACCTACTACCCAGGCATGGGCTAACACGACGAACGAAC  
TACTACTACTAAGCATAAAAAACAGGAAAAAAATGGAAAACTTAAAAAATGGATCATACAACCG  
30 AACGCAAACGACCTACGACGATCGATCTCACTTCCCCGTCTTTTTCATCCTAAGCAATAGAACGA  
TGGTAGAAAAGGAAGATAAAGATGGAGAGAAAGTCACGTGTATCAATGACGACGACTACCAAAA  
CTGAAGACGTAACACATGTTCCCCAGCGAGCGGTAACCTGTTCTGTTCTGACACCTTCCGCTCGA  
CAATGTACCTTTTAAAAACATACAAATTAGAAGTCGTCTTCACTACCTTCAACCAATCCAGCCAC  
TTTGGTATATACTTTTCATAGAATCCTTCTGAGCGCAAGGACCCTATTGAAATTCAGTGTTATTTT  
35 GTAACGCGACCAAAATGCCTAGCTGAATGTTGTTGAACGAGTTATGTACATCAAAAGATTGAATA  
AAACAAAAA

## Figure 2

*Anopheles gambiae* arrestin 1 amino acid sequence (SEQ ID NO: 2)

5

MVYNFKVFKKCAPNGKVTLYMGKRDFVDHVSGVEPIDGIVVLDDEYIRDNRKVFGQIVCSFRYGR  
EEDEV MGLNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKKLGSNAIPFTFNISP NAPSSVTLQQG  
EDDNGDPCGVSYVVKIFAGESETDRTHRRSTVTLGIRKIQFAPTKQGQQPCTLVRKDFMLSPGELE  
LEVTLDKQLYLHGERIGVNICIRNNSNKMVKKIKAMVQQGV DVVLFQNGSYRNTVASLETSEGCP  
10 QPGSSLQKVMYLTPLLSSNKQRRGIALDGQIKRQDQCLASTTLLAQPDQRDAFGVIISYAVKVKLFL  
GALGGELSAELPFVLMHPKPGTKAKVIHADSQADVETFRQDTIDQQASVDFE

Figure 3a

*Anopheles gambiae* odorant receptor 1 genomic sequence (SEQ ID NO: 9)

Features:

- 1) Presumed Untranslated 5' and 3' regions are underlined.
  - 2) Potential TATA box transcription initiation signal is double underlined.
  - 3) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
  - 4) Introns are tentatively assigned and are shown in lower case.
- Exons are highlighted.

AGCTTTGTTTCATTTATGTTGAAATCTAGCCCATTTTGTATAGTGCTGAACGACGAAGAACATACGAAAGTACCTCGT  
CCGAACACTATCAACATTAATTATACCAAGCTAGAAGAAGATATTTATAGTCAAGCCTCAACATCATAGGAACTTT  
AGCAAAACCATTTAATTTACATGATGATAAGTCCACCTCTTACCCAGCACAGGTTTGAGAAGGACGAAAGTATCT  
TTACGATAATATTACTCTAAGGTAGTTTTTGAATAAAATAAAAATTTACGTGCAAGTGGTGGCATCGGACATCATTC  
GAAAGAATCTACTAAGTCATACACACACCCAAAGACGACGCGTAGTTTCATCTAGAAAAACGGGTGAGCTCCATC  
GAACACGTGAGGACATAACTGCGACATGCGTATGGTCACTTAGTGCCAACACTGGTTCAGGGGACTACCTT  
CCGAAGCAGTAGAACCTAATGTATTGGAAATTTATAGGACATACTGCAACATGCATATGGCTAGTTCGCTGGTACC  
AACGATGGCACCAGGACACTATCTGCGGCCCTTGTAAATCACTGTAAATCTATACAAAAACGGCTTTACCCATACT  
TTATCACAAAAACGGCAGGTGAGGGCTGGATTGCTTCAAAGCATTAGAAATATATAATTTCAAAGTCCATAATCTCC  
TTAAAGATAGACAaCAGTAGAGAACACATTTAGTGCTCTTTTCGTTTCGAGTTAGTTGCCTTCTCAAGTAAGCGTTT  
AATGCTCAATTGTTGTAGATTGCTGGATGACTCTCGCTACGTGCTATAGTGGTCAATACTTCCAATTAGATTTCAT  
AATTAGTTTCCAATTGTCCACGGAAAACCCaCAAAAGAAAAAAACTTGTATCTAGGGTGGAAATTTTTCGAGAACA  
ATTGGACACTTCATATGAAAAAGGACAGCTTTTTCAAATGTTAAATAAACACCGTTGGATCCTTTgttgatttca  
atttctccaaattctgcagaataattctgcaattttacaaaactgctcaaccaccaataattccaattaatcatctg  
aacattttaaactgataaattagatgagtaattgcttcgctcatcacctaagaaatcgattagtttgataaaaaagaa  
caaattgaaatacaataaagtcctgaattttattcgaataaacggttgaaactcatttatttcaaaaacctttgaga  
aattcctcggttgaaaattgggtctcctatagttctgctaacggggccacttcaaaagcaagaactaacaataatcataat  
tatgggtgcaagtaactatcagtagcagtaatcgccattaaaaacttttctcaatttgcggtcgttaccgggctaaa  
tacagagcagagtaacgggaagtgatcaacgtcgctattagtagataacgaggaacgcctccgaagggtgtgtgaagg  
accttttcaaattgaaaccaagtactgtttccagttttaaattggatagttataaaatgagcgttcaacgatcggg  
catcatttgagtttcatcttcgaggagaaatagatcagtgccactgtttaaccgaaagtaataagctgaacaaact  
gaacccacggtgggatgcgtacgatcgacgggattcgttctggttgagttgctttgtttgaaatatttagGCCTAT  
GGCCACCGGAAGATACGGATCAGGCAACGCGGAACCGGTACATCGGTACGGTTGGGCTTTGCGGATCATGTTCTA  
CATCTGTACGGCTCTAACGCAAGCCCTATACTTCAAGgATGTGAAGGATATTAATgtgagttctctagtttagctattag  
tgttccacctgtccataatctgtcttttattgggttagGACATCGCAAAATGCATTGTTTCGTGCTTATCACTCAAGTGA  
CGTTGATCTACAAGCTGGAAAAGTTTAACTACAACATCGCACGGATTACGGCTTGTCTGCGCAAGCTTAACTGCACA  
CTGTATCAGCCGAAACAGCGCGGAAGAATTACgtaagcctgctgggaaatatgactaaaaagagtgctaacaacga  
ctctcctccaaatgtagCCCCGTTTTACAATCGATGAGTGGAGTGTTTTGGCTGATGATCTTTCTCATGTTTGTGGC  
TATCTTCAACCATCATCATGTGGGTATGTGCGCCAGCCTTCGACAATGAACGTGGTCTGCCCGTCCGGCCTGGTTCC  
CGGTGGACTATACCATTCGGACATAGTGTACGGTGACTGTTCTCTGATCAAACCATTTGGAATCGTCATGAGCGCA  
ACGTACAACCTTCTCGACGGATACCATGTTTTCCGGCTTGATGCTACACATAAATGGACAAATTGTGCGGCTTGGTAG  
TATGGTTAAAAAGgtgagttacggcgactacttgctccagtaaggacagggagtttggttccggttatgatattatt  
ttatcagCTTGGACATGACGTCCCTCCCGAACGCCAATTGGTTCGCAACGGATGCGGAATGGAAACAGATGCGAAAGC  
GCATCGACCATCACTCCAAAGTGTACGGTACGATGTACGCTAAAGTAACGGAGTGTGTGCTGTTTACAAGGACATC  
TTAAGgtacgaattgggccaattaattgtgtcatttaaaaagccttgacccaacttttcacagcttcggcgatgaagt  
gcaggacattttccaagGATCTATCTTCGCGCAAGTATGCGCGTCTGTAATTATCATTGTATGACACTGCTGCAAG  
TACCGGGGGCGATCTTACGATgGCCGATCTGCTGGGCTGTGGGTCTATTGCTAGTAAaGACATCGCAAGTGTTTA  
TTTTCTGTTACGTACGGAATCAAAATCTCCTATACGgtaggttgacacgtagaggaattaaatgtttgggaagaata  
tcaataccaaatagtagatgttttcggttacagACGGATAAATTTACAGACTTGTGTGGGTTTTCCAACACTACTTCAAG  
TTTGATAAGCGTACGACCAAGCAATGATATTTTTTCTGCAAAATgtgagatagcgggtgtatttgtgcagtcagtaca

ttaaatacgttctctattttcagCACTCTTAAAGATGTTACATCAAGGTGGGAAGTCTCTTGAAGCTTACGCTAAAT  
 CTTACACACATTTTGCAGgtatgtaattatgctgtggtatttagcttgaaataagctacaaactttgaaagtaattt  
 caatctgtttttagattATGAAGCTATCGTACTCCTATCTGGCCGTACTTCAGAGCATGGAATCAGAGTAATGGEG  
 5 TAAATATCCCTAA TGTTGAAATTATATTTGTTAGATTTATTGCATAAAGTAaTaTTTAATTTTATACATCAAACGT  
 AAGCCCGCtaGTTTTCAATTAGCCTTTCCAAAATTTATCAAATTGATTTTCAATTGATTGCAGAGTTTCAGGAATT  
 TAATCTGATAGGATATCTTGTATCCAATAGAGGTGTGGAAGCGTTCCAAGCCATTCGTTTGATAGTTTATAGCA  
 CCGTCGAGCAGTTGATCGCTGTGATCGCTAGGCGCACCTGATTTTATCTTTATCTCGCACCTGTTATGGCAAGGGCG  
 CTTTTACACGTTTTACACAATATAATGCACATGTATAATGCATTCTTACTTTAGCATTTTTGTACATATAATACC  
 10 AAAATTATGCATTTTTATTCTCACGCAACGATTAGAGGATGACTTcACAAAGGTCCATCTAGTGGTAGGAGGTATAC  
 AATTATACCTCTCAAAATCTCACAGCAtAATGAGAAACAAAAGGATACCAAGCATACCCTTTTTTTTACTTGACAATT  
 TCATTTGATTTATGTAATAAAGCACTGCaCGTCGACTTCCTAAAA

Figure 3a continued

**Figure 3b**

*Anopheles gambiae* odorant receptor 1 amino acid sequence (SEQ ID NO: 4)

5 MKKDSFFKMLNKHRLCLWPPEDTDQATRNRYIAYGWALRIMFLHLYALTQALYFKDVKDIND  
IANALFVLMTQVTLIYKLEKFNYNIARIQACLRKLNCTLYHPKQREEFSPVLQSMGVFWLMIFLM  
FVAIFTIIMWVMSPAFDNERRLPVPAWFPVDYHHSDIVYGVLFLYQTIGIVMSATYNFSTDTMFSG  
10 LMLHINGQIVRLGSMVKKLGHDPVPERQLVATDAEWKEMRKRIDHHSKVYGTMYAKVTECVLF  
HKDILRIYLRASMRVCNYHLYDTAATTGGDVTMADLLGCGVYLLVKTSQVFIFCYVGNEISYTDKF  
TEFVGFSNYFKFDKRTSQAMIFFLQMTLKDVHIKVGSVLKVTNLHTFLQIMKLSYSLAVLQSM  
ESEZ

15

bioRxiv preprint doi: <https://doi.org/10.1101/014340>; this version posted October 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



2042105-02402

5  
10  
15  
20  
25  
30

ttccagtaatccataataaaaaataataaaagtaaataaatagtaaataattccagtaactgtagtaatac  
acaataatctctaagaattaaaattgcatttttgtaatgaaatatgttgattgttcgaatagttcagaaaaacttaaa  
aatgcctcagcattaaacagttttgaggttggttcagggcatttagtttagatatttttagtatttttaaagcatttggt  
ttcattactacaaaaaagcaaatttatgagtgaattactttcagttcttctaaacgcctatgtgatatgcaattacat  
aacaatagctctcttttttattgcatttttcttagtaatcctaaatccaatctcttcttccctcttgagATTAAA  
GTGGGCAACGTGTACCCGATGACGTTGGAAATGTTTCAAAAATTGCTCAACGTGTCTACTCCTATTTACACTGCT  
GCGCCGAGTGTACAACATAACTTAACCGGTAAACAAACAAAAATCCCCTCATCACTATGCAAAGACAGCAAGCAGCC  
GATCATCAAACACCATTAGCAGCCACAAAGTTACCAGCCGCTTATCCACGGGATTGTTGGTGGAAAGTTATTGCACTG  
AAGCTCTTTTACCCCAAATTTTCATGGAGGTTCCCTCTCAACCAACCCATTGAAGCGAATAAAAGTATCAGCAACCAG  
GCGACGGTGAAAAACGCTGCATTATTGTGCTTGCTTCAGCATTCCAGCGAATGACTCTTAACTTTTCATTCAAA  
AGTCGCGATGCTCACGATACGGAGCGGTGTGTTGTTCGATCCGCCGAGTGCCTCGCAAGCCGGTGATGTTGCCGGT  
GGAAATGCACAGATCGACACAGCGATAGATAATCGTTTGTTCGCGTAAATGGGAGGGAAAAAGTAAGCTGCCAGCT  
ACTTCATTTCCATGTTAATTGAACTCAAGCCAACGAACATGCAGAACCCGGTTGGTTGTGTGTCTCCGCTCCGGGA  
AAGGTCTCTGCTCCGGGGCATGGATTCTTTCCCCCTCCGGGTGGTTGGGGGTATTGTTTAGGTTTTTATTTTACAAA  
TTCATATCCTTCCGCTTCCGCATCAGCCGACCCGGTGGGTGCGCCAGACAGATGTGCGGCGGGCAACAAAACATATGC  
ACGAACATGGCCAACAAACACAGCTTCTATCTCATCTCTGTGTGCGCACTGTCTCGCTTTCCCGCTGCGTTGCTTGTA  
GTACTATCATTGTTTTAGTCCACGGGTTTACTTCTAATTCCATTGCACCACGCAAAAAGGCTCATCCTTTGCTCGTT  
CCGGTTGCAACTTCGACAAGCGCATGGTTGGGATACGAACAAAAACCAACTACTCCACCCACTACTACTACTG  
CCACCACCACTAACAACACTACACTTGGTTGGGAGCTTGCAGACCCACAAGCAACAACGATACAAGCTAGCTAGCT  
GCTGTGTGCGCTCGAGTCAGCCGACGGTACAAGGTTTAACCGGTACAAGCAACTCCCGGACCGATCCCAAACTCTG  
ACAAGGCACGGGGCCGCATCCGGCAGTACGGTCGGAAAAACATGGAAATGTTTAAATTAAACTGTAATTGTCAATCGC  
TGCTACAAGTTGTGACACAGGGAGAGAGAGACAGAGCGCGCCCGATGGTGATGGTGTAAGATAGATACAGGAA  
AAGAGCGAGAAACATTGGTACGATTGGTGTGGTTAGCAAATTTGATTTCCTGATTTTGAGTGCAAATTTAATGC  
ATCGAAAATTTGCCATTGAGGTAAAGTTGCTCGTGACGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTAT  
CGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCTTTAGTGGA

Figure 4a continued

**Figure 4b**

*Anopheles gambiae* odorant receptor 2 amino acid sequence (SEQ ID NO: 6)

5 MLIEECPIIGVNVRVWLFWSYLRRPRLSRFLVGCIPVAVLNVFQFLKLYSSWGDMSELIINGYFTV  
LYFNLVLRTSFLVINRRKFETFFEGVAAEYALLEKNDDIRPVLERYTRRGRMLSISNLWLGFISA  
CFVTYPLFVPGRGLPYGVVTIPGVDVLATPTYQVVFVLQVYLTFPACCMYIPFTSFYATCTLFALVQI  
AALKQRLGRLGRHSGTMASGTGHSAGTLFAELKECLKYHKQIIQYVHDLNSLVTHLCLLEFLSFGM  
10 MLCALLFLLSISNQLAQMIMIGSYIFMILSQMFAFYWHANEVLEASLGIGDAIYNGAWPDEFPIR  
KRLILIIARAQPTDGGKIKVGNVYPMTLEMFQKLLNVSYSYFTLLRRVYN

15

[illegible]



Figure 5a

*Anopheles gambiae* odorant receptor 3 genomic sequence (SEQ ID NO: 11)

Features:

- 1) Presumed Untranslated 5' and 3' regions are underlined.
- 2) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
- 3) Introns are tentatively assigned and are shown in lower case.
- 4) Exons are highlighted.

AAGCAGAACACATCAAGAAGCAATTAGGTGTGTCTGACGTTAGCAAGTAGTTCGCGAGGAGGAATAAAATAGATGCC  
TTOTGAGCGGCTTCGTCTCATTACTTCCTTCGGAACCTCTCAAGAGAAACGCACGATGGTACTGCCAAAAATTAAAGG  
ATGAAACAGCAGTGATGCCGTTTCTGCTGCAAATTCAAAACATTGCCCGACTGTGGGGTGACCGTTGCCAGCGGTAC  
CGTTTTTATCTCATCTTTTCTTCTTCTCGCGCATGGTGGTTCTACCCAAAGTGCTGTTCGGTTATCCAGATCTCCA  
GGTTTGGCGTACCGCGCAGCGCGGAGCTGATGTTCCAATCGAACGCATTCTTCGGCATGCTAATGTTTTCTTTCAAG  
CGGCAAACTACGAGCGAATGGTGCATCAGCTGCAGGATCTGGCAGCTCTAGgtgagtatgcagccaatcgattgttc  
caaaccttcgcaacatccttcgtaacactgctacacttttcagTCCCTCCAAGACCTACCCACAGAGCTGGGAGAGTAC  
CTGATCTCAGTGAAACCGACGGTTCGATCGGTTCTCCAAAATTTACTGCTGCTGTCACTTTTCCATGGCAACGTTCTT  
TTGGTTTCATGCCGCTCTGGACGAGCTATTTCGGCTACTTTGCTGTGCGCAACAGCAGGAACCGGTGAGACAGTGT  
TGCAGCTCGAGGAGAGCTGTACTTCTGAAATTCGGACTTCGATGGCGCACTATACGTTTTATGTGGCCATTATG  
TGGCCACGATCTATACGCTCGGGTTTACCGGTGGCACAAAGCTGCTGACCATTTTCAGCAATGTTAAGTACTGTTTC  
GGCCATGCTGAAGCTCGTTGCACTCCGAATCCACTGTCTAGCGAGAGTAGCGCAAGACCGAGCGGAAAAGGAGCTGA  
AGGAGATTATTTCCATGCATCAGCGGCTACTCAAgtgaagtaaattcaaattgaaagttttgcaggggaataaacttgag  
tgtgtctgaccgctgcacatcctagTGGCTGTTCCTGCTGGAGACGACATTCCGCTGGGTATTTTCTGTGCAGTTTC  
ATTCACTGTACAAAGATCTGGTGCAGTCTCATCTCTACATAGCGGTGACGgtaatagcattttcgtcatttcgtta  
gccttattcaatccatttttgtgaacgtgaattttccccagCGGTTCACTCGACGGTAGCGAATGTATGTGTCCAG  
ATCAATTTGGTGACCGTGCAAACTTACGGCTACGGCTACTTCGGAACAGATCTAACACGGAGGTGCTTTGGgtacc  
ctttggatgaagcttcaaaaagtaattccaaattctgttttcgatttttcccttttccactagAGCTATGGCGTTG  
CGGTGGCCATTTACGATAGCGAGTGGTACAAAGTTTCCATTTGATGCGCGGCAAACTTCGACTGCTACTGCAACGA  
TCCCAAAAACCGGTGGCGTAACGGCGGGAAGTTTCGCTTCGTCAATGTGGCCCAAGTTTGGCAAGgtaacattaat  
tacagtttgaaaattctgaagaatgcactcttacttgcttacttggtgttccagATGCTCAAGATGTCTATTCTATT  
TTAGCTTAGTCAAGGAGCAGTTTATGAGCTGCTGTTTCCACCTGGAAATGGCCTTTTTCGCACTGTCTTCTGT  
TTGTTGGACGCACGCAGCACCAGAGAGCGCCCTGCACGCACTGACGTATTTTGGCTACTTTGACGTTTGCACCTTTG  
ACAGCTGAAGGACAGGGTACAATTTTTTGCTGCTGTTATTACGCGCAGCGCATTGGATACGAAAACATTGGCCACAAG  
TTCTACGATTTTAGCGTTTATTTACTGTTTCGTAGCAGCTTTTTCCaCAATAAACACACACAATAACGTACCGACAG  
TATTCTTTTCATTGTAGGATAGAGAAGCCCGGCCAGCAGCCAAAACGCGCCGCAAAACGAAAGGCGGCACCACCG  
GGGAAAAACACGGGAGCAAAACGAGAACAGAACGCAGTAACAACAAAACCGCCGGAACAACAACGGTGCCGGAA  
ACGA

## Figure 5b

*Anopheles gambiae* odorant receptor 3 amino acid sequence (SEQ ID NO: 8)

5 MPSERLRLITSFGTPQDKRTMVLPKLKDETAVMPFLLQIQTIAGLWGDRSQRYRFYLIFSIFCAMV  
VLPKVLFGYPDLEVAVRGTAELMFESNAFFGMLMFQFQRDNYERLVHQLQDLAALVLQDLPTL  
GEYLISVNRRVDRFSKIYCCCHFSMATFFWFMPVWTTYSAYFAVRNSTEPVEHVLHLEEELYFLN  
10 IRTSMAHYTTFYVAIMWPTIYTLGFTGGTKLLTIFSNVKYCSAMLKLVALRIHCLARVAQDRAEKEL  
NEIISMHQVRLNCVFLLETTFRWVFFVQFIQCTMIWCSLILYIAVTGFSSTVANVCVQIILVTVETY  
GYGYFGTDLTTEVLWSYGVALAIYDSEWYKFSISMRRKLRLLLQRSQKPLGVTAGKFRFVNVAQF  
GKMLKMSYSFYVVLKEQF

15

204210 504500

Figure 6a

*Anopheles gambiae* odorant receptor 4 genomic sequence (SEQ ID NO: 12)

5

### Features:

- 1) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
- 2) Introns are tentatively assigned and are shown in lower case.

10 GGGGAAC TCCCCACCCGACCGACGACGACGAAAGCTAACGATGTGCAATTGAATAGTCATTAGT  
AGCGTTTTTGTCTGCGAAACGAACTAACCCCTTTGACTTTTTTAAGTTCAC TACGGTGAGGACAAAA  
TCAATAAATTAAATCGAGACCGTTGATGAGCAAAAAGAAAAAAATATTTTACTGATTTTCATTT  
CGTTCCATCGACTACATAATCATAATTATATGCCACATTTTATTATAAGTTTTTGTATCATTTTTTA  
AACAAACAAAAAATGCATCCTTTTGAATATTAGTCAGGTTGTATCAACAATGAAGTTTGAAGTGT  
15 TTCAAAAATATTCTCCCGGACACGGTCTTATCCTTCGTGCTAAGGCTTTTGCATATCGTGCGGC  
ATGAATGGGGCAGGATTTTCGGTCGCGAATTCGAGTTGGTGGCATTTTTCTGTTCTATTTAATCTT  
TCTTGTAATACCGCCACTAACGGGCGGGTACACCGATGGTACCAGCGTGTACGCACCAGTGTG  
GAATTCCTGTTTAATTGCAATATTTACGGCGGCAGTATGTTCTTTGCCTACGATGTGGCCACTTT  
CCAAGCGTTCATCCAGGAACTGAAGAGCCTTTTCGGTTTTTGgtaatatattaattaataaaattgcgtttattgcat  
20 catcattgttttctctttgcagTATGCTCACATTCGTACAGACTAAAGTATAAGCTGACCCGGTTCAACCGTC  
GAGCGGATATTATCGCCAAAGTGCAAACGACCTGCATGGGTGCTGTAACGCTTTTTCTACTGGAT  
TGCACCGATAACCTTCCATCTGTGCGCACTACTACAGGTCGACCAATTCACCCGAACCCGTCGCG  
TTTGTGCAACATTTAGAGGTGAAGTTCTATTGGCTCGAGAATCGCACCTCAGTCGAGGACTACAT  
AACCTTCGTGCTGATCATGCTACCCGTCGTGGTTATGTGTGGTTACGTATGCAATTTGAAGGTGA  
25 TGACCATCTGCTGCAGCATTGGACACTGTACACTGTACACCAGGATGACTATAGAGATGGTAGA  
GCAGTTGGAAGCATGGCATCAGCGGAACGAACTGCCAGCGCCATACGCAACGTGGGGCAGAT  
GCACAGTGGTTTACTGAAATGCATTAGGCTTTTGAACACGTC AATCCGATCGATGCTGATGCTGC  
AGTGGTTGACCTGCGTGTTAAACTGGAGCATTTCTCTCATCTATCTAACGAACGTGgtagttttgtctt  
gtttggaaatccaaaaacaaaagatggctataattgaactttctattacagGGCATCTCGCTACAATCGGTTACCGTGGT  
30 GGTAATGTTTTTTCTTGCCACTGCGGAAACTTTCTCTGTATTGTTTACTTGGGACGCGGCTTGCGA  
CACAACAGCAGCTGCTGGAGCACGCACTCTATGCTACACGGTGGTACAAC TACCCAATAGCCTT  
TCGCAGCAGCATTAGGATGATGTTGAGACAGTCGCAAAGGCATGCACACATAACGGTGGGGAAG  
TTTTTTCGCGTTAATTTGGAAGAATTTAGCAGGATTGTCAACTTATCCTACTCTGCTTACGTCGT  
ACTTAAGGATGTAATAAAGATGGATGTACAGTGAATGTTTTTTTTTTTTTGGCTTGGCAACGAATGA  
35 AGTTTTCCGAATCTATATTAGATCTAGAATTTAATCTAGATGTCATAATATGATCTTGGCCATGA  
CCGGTTCTCTGGTTTTTGGAAACCAATTCTCAAAACAATTTTGAAGTTAGGGCGAGGCATGAAATGTC  
CCAAGAACCTATCCAAGTTCTGGAACTACATATTACCGAATCTATCCCATTTATTGCCTCGGAACT  
GGTTTGGTGCTAAATATTTGTCCAAATGTTGGTCCTGGACCTATCCAGACAAAGATCTTCAATTA  
TTCTTACCACTGGAAC TGATTAATTGATGTAGGAAGTCATGGAGGTGTTTCAGGGAGAATTTAAA  
40 CACTAATGTTCCAAC TCAATTTTCAAGGGCAATTCTATTTTTTATATGCCCTACGATTGATAC  
GTATGTATTACTCCATTTCTTGACTTTTGTCTTATTCTTGCTGCTGATTGGACGTGAAATGTTGA  
GAAAAAGATTCTTATTTATGAGTGATACAGAGCCTTTAAATACTCCTACGTTGTTTGCTATTTAA  
GTATGGCCAGGCTAATCACAATCGCTACTAATGAACAGAATCTCTTCTAATTAACCCCTTTTCGAT  
TGATAGTGTCAATGTCAATGTCGAGATAATTGAACTGCAAACgATACCTACCTTAAACGGAGCAG  
45 AACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTTCGCGAGGAGGAATAAAAT  
AG

50

### Figure 6b

*Anopheles gambiae* odorant receptor 4 amino acid sequence (SEQ ID NO: 14)

5

10

MKFELFQKYSSPDTVLSFVLRLLHIVGMNGAGFRSRIRVGGIFLFYLI FLVIPPLTGGYTDGHQVRV  
 TSVEFLFN CNIYGGSMFFAYDVATFQAFIQELKSLSVLVCSHSYRLKYKLTRFNRRADIIAKVQTT  
 MGA VTLFYWIAPIPSICAHYYRSTNSTEPVRFVQHLEVKFYWLENRTSVEDYITFVLI MLPPVVMC  
 GYVCNLKVM TICCSIGHCTLYTRMTIEMVEQLESMA SAERTASAIRNVGQMHSGLLK CIRLLNTSI  
 RSMLMLQWLTCVLNWSISLIYLTNVGISLQSVTVVVMFFLATAETFLYCLLGTRLATQQQLLEHAL  
 YATRWYNYPIAFRSSIRMMLRQSQRHAHITVGKFFRVNLEEF SRIVNLSYSAYVVLKDVIMDVQ  
 NVSYSYFTLLRRVYN

15

000	00
001	01
002	02
003	03
004	04
005	05
006	06
007	07
008	08
009	09
010	10
011	11
012	12
013	13
014	14
015	15
016	16
017	17
018	18
019	19
020	20
021	21
022	22
023	23
024	24
025	25
026	26
027	27
028	28
029	29
030	30
031	31
032	32
033	33
034	34
035	35
036	36
037	37
038	38
039	39
040	40
041	41
042	42
043	43
044	44
045	45
046	46
047	47
048	48
049	49
050	50
051	51
052	52
053	53
054	54
055	55
056	56
057	57
058	58
059	59
060	60
061	61
062	62
063	63
064	64
065	65
066	66
067	67
068	68
069	69
070	70
071	71
072	72
073	73
074	74
075	75
076	76
077	77
078	78
079	79
080	80
081	81
082	82
083	83
084	84
085	85
086	86
087	87
088	88
089	89
090	90
091	91
092	92
093	93
094	94
095	95
096	96
097	97
098	98
099	99

Figure 7

**ANOPHELES GAMBIAE**

**Preferred DNA Codons**

Amino Acids			Preferred Codons							
Alanine	Ala	A	GCC	GCG	GCT	GCA				
Cysteine	Cys	C	TGC	TGT						
Aspartic acid	Asp	D	GAC	GAT						
Glutamic acid	Glu	E	GAG	GAA						
Phenylalanine	Phe	F	TTC	TTT						
Glycine	Gly	G	GGC	GGT	GGA	GGG				
Histidine	His	H	CAC	CAT						
Isoleucine	Ile	I	ATC	ATT	ATA					
Lysine	Lys	K	AAG	AAA						
Leucine	Leu	L	CTG	CTC	TTG	CTT	CTA	TTA		
Methionine	Met	M	ATG							
Asparagine	Asn	N	AAC	AAT						
Proline	Pro	P	CCG	CCC	CCA	CCT				
Glutamine	Gln	Q	CAG	CAA						
Arginine	Arg	R	CGC	CGG	CGT	CGA	AGA	AGG		
Serine	Ser	S	TCG	AGC	TCC	AGT	TCT	TCA		
Threonine	Thr	T	ACG	ACC	ACT	ACA				
Valine	Val	V	GTG	GTC	GTT	GTA				
Tryptophan	Trp	W	TGG							
Tyrosine	Tyr	Y	TAC	TAT						

[http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Anopheles+gambiae+\[gbinv\]](http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Anopheles+gambiae+[gbinv])

**Figure 8**

Name	SEQ ID NO	FIG. Reference
Arrestin 1 (cDNA)	SEQ ID NO: 1	Figure 1
Arrestin 1 (polypeptide)	SEQ ID NO: 2	Figure 2
Odorant Receptor 1 (cDNA)	SEQ ID NO: 3	—
Odorant Receptor 1 (polypeptide)	SEQ ID NO: 4	Figure 3b
Odorant Receptor 2 (cDNA)	SEQ ID NO: 5	—
Odorant Receptor 2 (polypeptide)	SEQ ID NO: 6	Figure 4b
Odorant Receptor 3 (cDNA)	SEQ ID NO: 7	—
Odorant Receptor 3 (polypeptide)	SEQ ID NO: 8	Figure 5b
Odorant Receptor 4 (cDNA)	SEQ ID NO: 13	—
Odorant Receptor 4 (polypeptide)	SEQ ID NO: 14	Figure 6b
Odorant Receptor 5 (cDNA)	SEQ ID NO: 15	—
Odorant Receptor 5 (polypeptide)	SEQ ID NO: 16	Figure 9b
Odorant Receptor 6 (cDNA)	SEQ ID NO: 17	—
Odorant Receptor 6 (polypeptide)	SEQ ID NO: 18	Figure 10b
Odorant Receptor 7 (cDNA)	SEQ ID NO: 19	—
Odorant Receptor 7 (polypeptide)	SEQ ID NO: 20	Figure 11b

Figure 9a

*Anopheles gambiae* odorant receptor 5 genomic sequence (SEQ ID NO: 21)

5

Predicted Exons: *ITALICIZED*, UNDERLINED AND **HIGHLIGHTED**.  
Introns: lowercase.

10 tctagacttgaacccatgacgggcattttattgagtcgttcgagttgacgactgtaccacgggaccaccggtttatcactatcactatt  
aattaattataatatgctttttagcgatcagcctaccgggtttgtttctctggatatcttaagttccatttgattatcaagatagaa  
caacaacttgtaccttaataatcattacgtacccttaataacgtgcatcaaggagttttcgcgaaagcaaaaaatccgattgtct  
gatgttgcttgattccatccgattcggttactggttctgcaaaatcgccaataatacggcaatgtccttatcgatgcttgaatcaacat  
cacattgtttgcatttcgtttttgcggtgcaaatatgttatttgcaagaaggcaaggtaatgtgcttaagagtaaatacaattcgctg  
15 tccattttttgcccaccagtggtccagaacccgtgccttttagtccctcgaatacatccgaccagtcagcaagcaagtgcac**ATGG**  
**TGCTACCGAAGCTGTCCGAACCGTACGCCGTGATGCCGCTTCTACTACGCCCTGCAGCG**  
**TTTCGTTGGGCTGTGGGGTGAACGACGCTATCGCTACAAGTTCGGGTGGCATTTTTA**  
**AGCTTCTGTCTGCTAGTAGTTATTCCGAAGGTTGCCTTCGGCTATCCAGATTTAGAGAC**  
**AATGGTTCCGCGGAACAGCTGAGCTGATTTTCGAATGGAACGTACTGTTTGGGATGTTG**  
20 **CTGTTTTCTCTCAAGCTAGACGACTATGATGATCTGGTGTACCGGTACAAGGACATATC**  
**AAAGATTG**gtgctgataatgattgataaaaggaacctttgagcaactcctatccctttcaag**CTTTCGGTAAGGAC**  
**GTTCCCTCGCAGATGGGCGACTATCTGGTACGCATCAATCATCGTATCGATCGGTTTTTC**  
**CAAGATCTACTGCTGCAGCCATCTGTGTTTGGCCATCTTCTACTGGGTGGCTCCTTCGT**  
**CCAGCACCTACCTAGCGTACCTGGGGGACGAAACAGATCCGTCCCGGTGGAACATGT**  
25 **GCTACACCTGGAGGAGGAGCTGTACTGGTTTCACACCCCGCTCTCGCTGGTAGATTAC**  
**TCCATATTCACCGCCATCATGCTGCCTACAATCTTTATGCTAGCGTACTTCGGTGGACT**  
**AAAGCTGCTAACCATCTTCAGCAACGTGAAGTACTGTTTCGGCAATGCTCAGGCTTGTG**  
**GCGATGAGAATCCAGTTCATGGACCGGCTGGACGAGCGCGAAGCGGAAAAGGAAGTGA**  
**TCGAAATCATCGTCATGCATCAGAAGGGCGCTAA**gtaaggtctgccggtatgttggtgtagaatacattt  
30 ctagctgctttcag**ATGTGTGGAGCTGTTGGAAATCATCTTTCGGTGGGTTTTTCTGGGACAG**  
**TTCATAAGTGCGTAATGATCTGGTGCAGCTTGCTTCTGTACGTCGCCGTTACG**gtaacta  
aaagcactgtagtgatctgtctgccacaccattcactgtgtgtcttgtttgtcactcttcccag**GGTCTCAGCACAAAAG**  
**CGGCAAACGTGGGTGTAAGTTTATACTGCTAACAGTGGAAACCTACGGATTCTGCTA**  
35 **CTTTGGCAGTGATCTTACCTCGGAGGCAAGTTGTTATTCGCTGA**gtttcagttacttttccgttccc  
tctaaccgtaccacttgtagcattgtttgagacagagcttgagcgtag**CACGTGCTGCGTACGGTAGCCTCTGG**  
**TATCGCCGTTTCGGTTTCGATTCAACGGAAGCTTCGAATGGTACTGCAGCGTGGCCAGA**  
**AACCGGTCCGCATCTCGGCTGGGAAGTTTTGCTTCGTCCGACATTGAGCAGTTTGGCAA**  
40 **TgtatggggagaccttcactgtggcaagaaagattttcttattaatgcatttttaatttacagATGGCAAAAACATCA**  
**TACTCGTTCTACATCGTTCTGAAGGATCAATTTTAA**aggggaactccccaccgaccagacgagga  
agctaacgatgtgcaattgaatagtcattagtagcggttttctcgcaaacgaactaacccttgacttttaagttcactacggtgag  
gacaaaaatcaataaataatcgagaccgtgatgagcaaaaagaaaaaataattttactgattttcatttcgttccatcgacta  
cataatcataattatatgccacattttattataagttttt

45

Figure 9b

*Anopheles gambiae* odorant receptor 5 amino acid sequence (SEQ ID NO: 16)

5 MVLPKLSEPYAVMPLLLRLQRFVGLWGERRYRYKFRLAFLSFCLLVVIPKVAFGYPDLE  
TMVRGTAELIFEWNVLFGLMLLSLKLDDYDDL VYRYKDISKIAFRKDVPSQMGDYL VRI  
NHRIDRFSKIYCCSHLCLAIFYWVAPSSSTYLAYLGARNRSVPVEHVLHLEEELYWFHTR  
VSLVDYSIFTAIMLPTIFMLAYFGGLKLLTIFSNVKYCSAMRLVAMRIQFMDRLDEREA  
10 EKELIEIIVMHQKALKCVELLEIIFRWVFLGQFIQCVMIWCSLVLYVAVTGLSTKAANVG  
VLFILLTVETYGFCYFGSDLTSEASCYSLTRAAYGSLWYRRSVSIQRKLRMVLQRAQKP  
VGISAGKFCFVDIEQFGNMAKTSYSFYIVLKDQF

15

[illegible]



Figure 10a

*Anopheles gambiae* odorant receptor 6 partial genomic sequence (SEQ ID NO: 22)

- 5 These are the predicted last three exons of another candidate *Anopheles gambiae* odorant receptor.

Predicted Exons: *ITALICIZED*, UNDERLINED AND **HIGHLIGHTED**.  
Introns: lowercase.

10

15  
20  
25

aacacccatcttatcggcaaaattagtagtattaccgtttgaaagcggcttccttcctggctgtttctcactctctctctctgtctctctta  
ttgatgccgtatgcgccgctgctataggctagTTATGCTTACCGGATGTTGCGATCGCGCACGTGCTTT  
TCCGCATACGCCAGTGCACACTTGATGGCGGTGGTGATGACGTCTGCTGCGCACCGTT  
TTCTGCTCGTGAGTCAGACCTTTTCATTTCTGCAATATCCTGTTTCTTTCCCGACCCC  
ACAGACGGTTAGACGGATATATGCTGGTAAAGTTTGTCTCTTCATGCTGTGCTTTCTG  
ATCGAGCTGCTGATGCTGTGTGCGTACGGTCAGGATATTGTGGAATCGgtaaggcaccagge  
ggtgatgagcgagtcgcgagtaattgaagcttttgccttttaaacaacacatcagagCCTTGGGGTGATTGATGCCGCT  
TACGGTTGCCAATGGTACCGGGAAGGGTCGGTGGCGTTCCATCGATCCGTGCTGCAAA  
TTATACACCGCAGCCAGCAGTCCGTCATACTGACCGCATGGAAAATTTGGGCCATCCAA  
ATGAGTACTTTTCACTCAGGgtgagttgccaattgattgcggttgcttaattttcagtaagagtgcgctctttcccttag  
ATCCTGCAAGCTTCCTGGTCCCTACTTTACCCTCCTGAAGACCGTCTACGGGAATAAgtaa  
gcgcgagagagagagagagagcagtagtgcgttcaccccttgatgaatcaatagatttctaatacatgaaccattgaaaaatgaatca  
acattttcgtagttgcacaatattgtaccattctatacagcttcaccacgaccaagcgtttgttgcacagaccacacacgtttcga  
caagccgcgtcacctgctggc

# Figure 10b

*Anopheles gambiae* odorant receptor 6 partial amino acid sequence  
(SEQ ID NO: 18)

5

LCLPDVAIAHVLFRIRQCTLDGGGDDVCCAPFSARESDFISCNILFLSRPHRRLDGYML  
VKFVLFMLCFLIELLMLCAYGEDIVESPWGDZCRLRLRMVPGRVGGVPSIRAANYTPQP  
AVRHTDRMENLAHPNEYFQSDPASFLVLLYPPEDRLRE

10

20420"50493007



1006405 01243

5  
10  
15  
20  
25  
30  
35  
40  
45

tgctacaacacattttatgcttcacagatttacttctgctgttttcgatggccagagcaacctcgcggtatgtcatgttctgctcctggg  
tgctgctagcctgcgagcagctgcaacacttgaaggtaggtacggtagcaaacgtggtgtctttacatccgctgcagcattatcct  
tategacgtgtagtgtaaaggtaaaagaggaagcgataaaaaagcaacattctctcacacctcgatctctctttatctctctct  
ctccatctctctcgggcagGGTATTATGCGATCGTTGATG  
GAGCTTTTCGGGCTCGCTGGACACCTACCGGGCCCAACTCTTCGCAACTGTTCCGAGCAA  
TTTCAGCCGGTTCCAAATCGGAGCTGATCATCAACGAAGgtatgtgaaacgtgtgctcgtggcagacg  
gactcaaagagagcataacacaatcccctggtagttcatttcaatgaccttaacactcggcaagctaagcgagacagtggggacag  
tgagaaagagagaacaagaaaaaaaaccatcatccgtacgacatcatcgctacgtaccggtatttcaggatgaggaaataaaac  
gctaggggaatgaaagtgcgacagaatgataaaacaatccccacccaggccccagcctggacgaacggatgtagtgtgcaagc  
gagcaaaaaaagtcaataaattgaagtttaaaaaatagattttcccgtccatccgtggtggagcgtaaagccccggcgacaactt  
cgagcacggcgaccgtgcacagtactgtgccacagttgtagggacggataagctccgttctttttatccttttttttgagatttgt  
ttcgcttcgcatcgtagacgagcttagtgccgtgttgccttaattgctatttattataaagcgttccaaatagaagatcggttctctc  
catttaatctatcgcgctgtacgctgaaactatgcactgtgctgtgaaacgtcaagctcgagcacgacgaatggcccaccgtacc  
acgcccgtggtgccc aaagcgcaacggaattgcatgttaacaaacctttgcctaccatccaatccgtgtgaaattgccgctctcttt  
ctctcttttgcgctttcggtgtatcgaacggtttgtccctttttttactttgtctcttgatctcttgctgtgctcactttcatctcatgtttgc  
ctgacgggtggtgggttttcgaaaaagagcgatttctctcgctgtgtgtgtgtgtgttttttaataaccgctccaggctggtgaacg  
ctgcaggaccgatcgagctagtttattatcagcttagtggttatccacccatgccccacatcacgtctgtggagagtgggggaag  
cttaagtccaatgtaatttaccgtgtttctgtcgcttcgtcaccttcttcgctgatggagattgggtcggttggcacgataaaagcccact  
gcacgttacggaccgagggaaaggtctttttgtaggcctagcaacggctcctcattaccgcatgggggtgtagctcagatggttagag  
cgctcgcttagcatgtgagaggtaccgggatcgataccggcatctccaacccacacaaaacgttttttaagaagatttttagggaa  
gatattaacgcggtacactgtgctcctctaagttggaagagtagatgagatgatgacaagggagaaggaacatgtgtacgtgtt  
gatagcaaacacacaaacaacaatatctctgataataatctgatgtgtgtgtgtgtgtattgtgttatgtcgtcttgcctct  
tgtccctctctctctgttcaactcctaaaaagaattgtttggagtcctctcagttcctcgtaaagatccttcgagattctcttctctt  
attatttattccacgagcctctgacataagtagccttccgcttatttcttctccttgacttgcagttccgtgtagagcgtcattttgag  
gtttacacatttccaccgacgctgattgttacattgtcatctacattgtttccgtttaccgttccgccccttttttaacgctaccaca  
gAAAAGGATCCGGACGTTAAGGACTTTGATCTGAGCGGCATCTACAGCTCGAAGGCGG  
ACTGGGGCGCCGAGTTCCGTGCGCCGTGACGCTGCAAACGTTCCGACGAGAATGGCAG  
GAACGGAAATCCGAACGGGCTTACCGGGAAGCAGGAAATGATGGTGCGCAGCGCCATC  
AAGTACTGGGTCCGAGCGGCACAAGCACGTTGTACGgtaggtatggttaatttctaaggtgtggtgtaaag  
cctccaggttccatgaaaaagggatactttaccacagtaagagttgttttgcgtgacttacattctttggagcattgtttggtgtgtg  
ctgaaaccggttgcaatatcgttttggaagaaattatgtgtaaagcgtattacaatctcattcctctgttaatctgtaccaattgtgtc  
agccccgaccgaaagcaggcctaattcgtaccagaaaaaccacaagctgtttgtaagcatcgatacggccgaagcttcaatccagc  
caaggcgccacctaactattgacgtgacttttgcacgttcacactctccctctccattctttctataaccaatcgctcgtcagccagcat  
cgcccgagtgaaagtttttattgaaacgatatacccgatcgattttccactaaacatgcttaaatcgtttcacaaagctccccaaa  
atcccattcaccaatccaccaatttgaagtcgctcgtctttgtgtccttgtgtttgtgtgtgtgtgtgagctggagacatgggggagt  
gagtaaccgaacaacctcttgccgtgttcacgatatacgaacgaccaagataagcatcccttttccctagccgatgtctccgata  
tctcgattccgcttccagcgaggcaaaagaaaaaggcgaactggctgacctcaccggggcgaggaaaaagcgtagggattacgtc  
gagcagcagagttgtgatttcttcttcttctgttccataaatcgctgacggtttccattaccgctcgaggagtgcacacacgtgaag  
ggaaagcgaaaacgttttagattccagcagcaacggcagcaccagaagcagcagcagcgcggaattgaatcatcctgacggat  
gagttgtctgggttttcgggtcggtggttacagcaccacaccatctgctgcagctaataacagctgtaaatctcgtagacatagactt  
gattttacaatattacacacacacttacacacacagctatagatttgcgttggcgatggctctgtacggcgtgacctacatgccgc  
gagccgtgttgcgtggttgcgatacggatcacgtccgattcgattcagcctcgctgttttgggtgaagatccttatcggtgacct  
ttcagtggtcgagagcgagggtcactatggcgctgtcagttggaaagctaggtcgtattcaaagggccattgtgccagtgttctt  
ttaagatagcgataagcttttgatcgaaatagtaaataacacattgtttcttttctattccaaactgttgcaacctcattattacg  
tttttgacggggtgtatagtaaattgcatactttaaggcgtgattttcaaatgtagcgttccgtatgcagaaacgcatggtattatgc  
aatttaacaatgctgcttcttaacattcaataacggcttattaaggaaacttttgcattgttttaacagcaaatagttagc  
tcagaacgatcacatttagtatcgttcaacaaagaactcttttaacacacaatttgaatgccattccctcgagaaagtttctgtc  
agtctcctctgcatcacagcaacaacaaacctgctcatgtttcctgctgttccctagctgtttgaacgttatttccgattcctgtgct



gtaccgcaccgcatccgtaccgataccggaacaaacggtgtgcgcgaaagaatccgctagcagccccactggcacgggtatttgctt  
ttggttctgtgtttttcttccactggtttgggtgcctgggcgaaggctagctcggtactttccggggccgcaatttctgcagccaag  
gcggcgtgctcgtggggccaaaagaat

5

Figure 11a continued

1005405-01407

# Figure 11b

*Anopheles gambiae* odorant receptor 7 amino acid sequence (SEQ ID NO: 20)

5 MVLIQFFAILGNLATNADDVNELTANTITTLFFTHSVTKFIYFAVNSENFYRTLAIWNQT  
NTHPLFAESDARYHSIALAKMRKLLVLVMATTVLSVVAWVTITFFGESVKTVLDKATN  
ETYTVDIPRLPIKSWYPWNAMSGPAYIFSFIYQVRWRNGIMRSLMELSASLDTYRPNSSQ  
10 LFRAISAGSKSELIINEEKDPDVKDFDLSGIYSSKADWGAQFRAPSTLQTFDENGNGNP  
NGLTRKQEMMVRS AIKYWVERHKKHVRLVSAIGDTYGPALLHMLTSTIKLTLLAYQA  
TKIDGVNVYGLTVIGYLCYALAQVFLFCIFGNRLIEESSVMKAAYSCHWYDGSEEAKT  
FVQIVCQQCQKAMTISGAKFFT VSLDLFASVLGAVVTYFMVLVQLK

15

204219"50495001